Combining Training and Quantized Feedback in Multi-Antenna Reciprocal Channels

Umer Salim, *Member, IEEE,* David Gesbert, *Fellow, IEEE,* and Dirk Slock, *Fellow, IEEE*

Abstract

The communication between a multiple-antenna transmitter and multiple receivers (users) with either a single or multiple-antenna each can be significantly enhanced by providing the channel state information at the transmitter (CSIT) of the users, as this allows for scheduling, beamforming and multiuser multiplexing gains. The traditional view on how to enable CSIT has been as follows: In time-division duplexed (TDD) systems, uplink (UL) and downlink (DL) channel reciprocity allows the use of a training sequence in the UL direction, which is exploited to obtain an UL channel estimate. This estimate is in turn recycled in the next downlink transmission. In frequency-division duplexed (FDD) systems, which lack the UL and DL reciprocity, the CSIT is provided via the use of a dedicated feedback link of limited capacity between the receivers and the transmitter. In this paper, we focus on TDD systems and put their classical approach in question. We show that the traditional TDD setup above fails to fully exploit the channel reciprocity in its true sense. In fact, we show that the system can benefit from a combined CSIT acquisition strategy mixing the use of limited feedback and that of a training sequence. This combining gives rise to a very interesting joint estimation and detection problem for which we propose two iterative algorithms. An outage rate based framework is also developed which gives the resource split between training and feedback. We demonstrate the potential of this hybrid combining in terms of the improved CSIT quality under a global training and feedback resource constraint.

Index Terms

Umer Salim is with Intel Mobile Communications, 2600 Route des Crêtes, 06560 Sophia Antipolis, France (email: umer.salim@intel.com). David Gesbert and Dirk Slock are with Mobile Communications Department of EURECOM, BP 193, F-06904 Sophia Antipolis France. (email: david.gesbert@eurecom.fr; dirk.slock@eurecom.fr). Some part of the material in this paper appears in [1] and was presented at the IEEE GLOBECOM 2009.

Manuscript received December 3, 2010; revised July 4, 2011; accepted November 8, 2011.

Broadcast channels, CSIT acquisition, MIMO systems, Quantized feedback, Random vector quantization, Reciprocal channels, Training.

I. INTRODUCTION

Multiple-antenna transmitters and receivers are instrumental to optimizing the performance of bandwidth and power limited wireless communication systems. In the downlink (DL), in particular, the communication between a multiple-antenna enabled base station (BS) and one or more users with either a single or multiple antenna each can be significantly enhanced through the use of scheduling, beamforming and power allocation algorithms, be it in single user or multi-user mode (spatial division multiplexing). To allow for beamforming and/or multi-user multiplexing capability, the BS transmitter must however be informed with the channel state information (CSI) of each of the served users [2] [3] [4], except when the number of users reaches an asymptotic (large) regime in which case random opportunistic beamforming scheme can be exploited [5], [6]. This has motivated the proposal of many techniques for providing the channel state information at the transmitter (CSIT) in an efficient manner. Proposals for how to provide CSIT roughly fall in two categories depending upon the chosen duplexing scheme for the considered wireless network. In the case of time-division duplex (TDD) systems, it was always assumed that CSIT should exploit the reciprocity of the uplink (UL) and DL channels, so as to avoid the use of any resource consuming feedback channel [7], [8], [9]. The way reciprocity is exploited in the current TDD systems, is through the use of a training sequence sent by the user on the UL, based on which the BS first builds an estimate of the UL channel which in turn serves as an estimate for the DL channel in the next DL transmission [7], [8], [9]. In frequency-division duplex (FDD) systems, UL and DL portions of the bandwidth are normally quite apart and hence the channel realizations can be safely assumed to be independent of each other. This lack of channel reciprocity motivates instead the use of a dedicated feedback link in which the user conveys the information, about the estimated DL channel, back to the BS. Recently, several interesting strategies have been proposed for how to best use a limited feedback channel and still provide the BS with exploitable CSIT (see [10], [11], [12] for feedback acquisition and [13], [14], [15] [16] for limited feedback based precoding and scheduling).

Although in the past, the balance has weighed in the favour of FDD systems when choosing a duplexing scheme (in part because of heavy legacy issues in voice oriented 2G networks and also because of interference management between UL and DL), current discussions in the standardization groups indicate an increasing level of interest for TDD for upcoming wireless data-access networks (e.g.WiMax, etc.). This interest mainly stems from the fact that TDD systems are extremely flexible in managing asymmetric

UL and DL traffic loads and secondly because they are seen as more efficient in providing the CSIT required by several MIMO DL schemes, thanks to channel reciprocity. In practical systems though, the use of channel reciprocity faces some limitations mainly due to the difference in transmit/receive RF electronics [17] and some calibration is required [17], [18] for reciprocity exploitation.

In this paper, we focus on the problem of CSIT acquisition in a TDD system. We take a step back and shed some critical light on the traditional approach above consisting in exploiting the channel reciprocity via the use of training sequences exclusively. In fact we show that this approach fails to fully exploit the channel reciprocity. The key shortcoming is as follows: when sending a training sequence in the UL of a traditional TDD system, the user allows the BS to estimate the channel by a classical channel estimator (it can be a least-square (LS) estimator or minimum mean square error (MMSE) based, see [19] for details). However, note that the user itself has the knowledge of the channel coefficients (obtained from DL pilots during the current frame or from the DL synchronization sequence or other control signals or even from the previous DL frames if the channel is correlated in time) but, regretfully, does not exploit that knowledge in order to facilitate the CSIT acquisition by the BS. Instead, it uses this knowledge only locally.

Interestingly, by contrast, in FDD systems, the user exploits its DL channel knowledge by quantizing the channel and sending the result over a dedicated feedback link (actually UL bandwidth is used for this feedback along with UL data transmission). In the FDD case, UL training is used by the BS solely for UL data detection as this UL training cannot give any direct information to the BS about the DL channel coefficients.

In this paper, we point out that in TDD systems there is a unique opportunity to combine both forms of CSIT acquisition. In doing so, we obtain a new CSIT acquisition scheme mixing the classical channel estimation using training with the quantized limited channel feedback of the same channel. This gives us a framework for fully utilizing the channel reciprocity in a TDD setup and it improves the classical trade-off between the CSIT quality and the amount of training/feedback resource used. We characterize the optimal CSIT acquisition structure under this novel framework. A novel hybrid CSIT acquisition setup is proposed which gives rise to a very interesting joint estimation and detection problem for which we propose two iterative algorithms. We further propose a sub-optimal outage rate based approach which helps us to optimize the fixed resource partitioning between training and quantized feedback phases. We adapt this optimization framework to use it with practical constellations like QPSK and 16-QAM. The results obtained confirm our intuition and clearly demonstrate the benefit of this hybrid (mix of training and quantized feedback) approach for upcoming TDD systems.

In previous work, Caire et al. studied the achievable rates for multi-user MIMO DL removing all the assumptions of channel state information at the receiver (CSIR) and CSIT for FDD systems in [20]. They gave transmission schemes incorporating all the necessary training and feedback stages and compared achievable rates either with analog feedback or with quantized feedback. The reference [21] studies the decay rate of the feedback distortion versus SNR with analog and digital quantized feedback for FDD systems. A very recent paper [22] studies combining the analog and digital feedback for FDD systems. Another recent reference [23] does a simulation based comparison of separation and non-separation based feedback schemes. All of these works fundamentally differ from our work as there is no channel reciprocity in FDD systems and hence there is no point in combining the UL training and the quantized feedback of the DL channel.

Some other contributions [7], [24], [25], [26] and [27] analyze the sum rate of TDD systems starting without any assumption of CSI but restrict the CSIT acquisition through training only. [8] does a comparison of TDD systems versus FDD systems in terms of CSIT acquisition accuracy. [28] studies the diversity-multiplexing trade-off [29] of two-way SIMO channels when TDD is the mode of operation. All of these references treat no-CSI TDD systems but all acquire CSIT through training only. According to authors' knowledge, there is no single contribution which exploits the combining of training and the quantized feedback in TDD systems, which we believe to be one of the major novelties of this work.

The paper is structured as follows: The system settings are given in section II along with the classical CSIT acquisition for FDD and TDD systems. The optimal CSIT acquisition strategy combining training and feedback is outlined in section III. Two iterative and one non-iterative algorithms for the joint estimation and detection have been proposed in section IV. The simplified outage-rate based framework to optimize the resource split appears in section V followed by its adaption for practical constellations in section VI. The simulation results have been provided in section VII, followed by the conclusions and the possible future extensions combined in section VIII.

Notation: $\mathbb{E}$ denotes statistical expectation. Lowercase letters represent scalars, boldface lowercase letters represent vectors, and boldface uppercase letters denote matrices. $\mathbf{A}^{\dagger}$ and $\mathbf{A}^{-1}$ denote the Hermitian and the inverse of matrix $\mathbf{A}$, respectively. For a vector $\mathbf{a}$, $||\mathbf{a}||$ and $\bar{\mathbf{a}}$ represent, respectively, its norm and unit-norm direction vector so that $\mathbf{a} = ||\mathbf{a}||\bar{\mathbf{a}}$. A Gaussian distributed vector $\mathbf{a}$ with mean $\mathbf{m_a}$ and covariance matrix $\mathbf{K_a}$ is represented as $\mathbf{a} \sim \mathcal{CN}\left(\mathbf{m_a}, \mathbf{K_a}\right)$. $\mathbf{I_M}$ represents the identity matrix of M dimensions.

II. SYSTEM SETTINGS AND PRELIMINARIES

We consider the two way TDD communication in a cell between a single BS, equipped with M antennas, and a single antenna mobile user. The channel $\mathbf{h} \in \mathbb{C}^{M\times 1}$ is assumed to be flat-fading with independent complex Gaussian zero-mean unit-variance entries, where $\mathbb{C}^{M\times 1}$ represents the M-dimensional complex space. A general TDD frame structure is shown in Fig. 1.

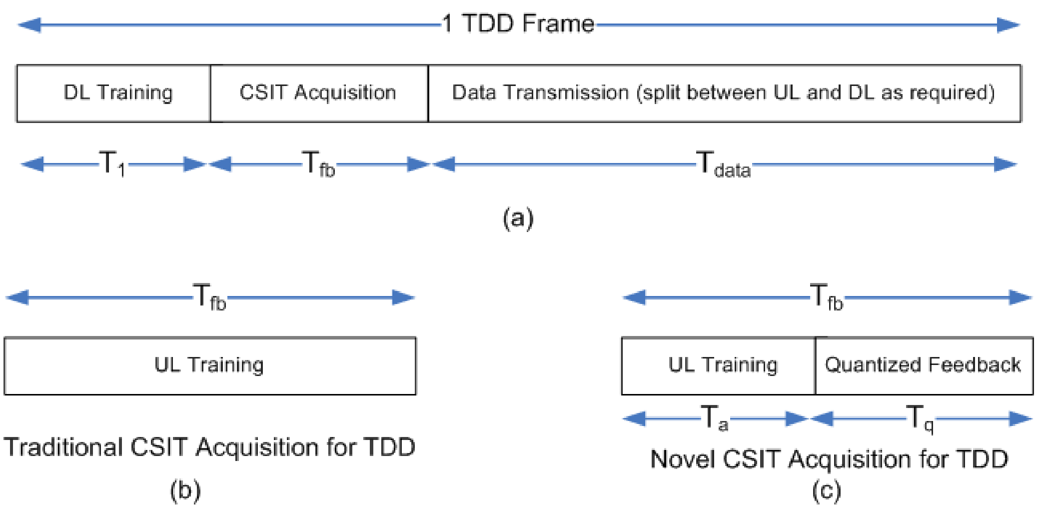


Fig. 1. (a) One TDD Frame; (b) Traditional CSIT Acquisition Setup for TDD systems; (c) Novel CSIT Acquisition Strategy for TDD systems: Total feedback length is divided between UL training and quantized feedback phases.

We assume that the channel realization stays constant for the duration of each frame. This implies that the channel coherence time is larger than or equal to the length of one frame. This channel model assumption is widely accepted in wireless systems [30], [31], entitled as block fading channel [32].

In TDD systems, UL and DL data transmissions are carried out in each single frame over the same frequency. Hence both the users and the BS need to have some reasonable channel knowledge for proper UL and DL operation. Fig. 1(a) shows that the TDD frame has been split in three phases:

- DL training: The first phase of the frame is reserved for DL training. In this phase, the BS will transmit global pilots which will enable all the users in the cell to estimate their corresponding channel realizations. In cellular systems, the users are always obliged to decode some low rate DL control information which requires the presence of DL pilots.
- CSIT acquisition: This phase is dedicated for CSIT acquisition at the BS. In traditional TDD wireless systems, the active users will send orthogonal training sequences in the UL direction and the

channel estimation carried out the BS based upon these pilot sequences furnishes CSIT [7] [8], as shown in Fig. 1(b). This CSIT is then employed for DL beamforming/precoding necessary for user multiplexing.

- Data transmission: The third phase is dedicated for the transmission of UL and DL information data. The split between UL and DL data portions can be carried out as required during this particular frame and can be modified later if desired.

We assume all the UL (as well as the DL) data symbols to be contiguous in each frame. This is commonly employed in practical TDD systems as well, for example in LTE TDD (consult table 4.2-2 in [33] for various UL-DL configurations in each TDD frame). If transmission is quickly switched between UL and DL, a guard interval is used (not shown in Fig. 1) as is employed inside the special subframe which is part of each LTE TDD frame [33].

The goal of this work is to provide a reliable estimate of the DL channel to the BS, which in turn can be used for beamforming/precoding purposes. It is true that the further optimization of the DL vs. UL time ratio could be of interest in the future. However in this paper, we focus on the situation where the UL to DL ratio is either imposed by the natural symmetry or the level of asymmetry of the traffic. Since a detailed model of traffic symmetry/asymmetry is beyond the scope of this paper, we leave this optimization aside assuming that the UL vs. DL split is governed by the traffic model. We would like to point out that the key idea of the paper about exploiting the reciprocity and the shared channel knowledge in TDD systems stays valid irrespective of the nature of the traffic.

The mean square error (MSE) of CSIT is selected to be the performance metric for CSIT reliability. It has been widely shown in literature that the DL throughput of a system with imperfect CSIT incurs a loss which is the product of the DL power and the MSE of CSIT [10], [20]. Hence the minimization of the MSE of CSIT is equivalent to the maximization of the system wide sum rate, the most commonly adopted system performance metric. As CSIT reliability directly translates into system performance, we can limit ourselves to the acquisition issue of the channel knowledge and its quality for a fixed acquisition resource and not about its use in MIMO transmission schemes. The similar strategy was adopted in [8] and [34] to investigate the system performance.

A. CSIR Acquisition at the Users

The BS transmits global pilots in the DL direction which are known sequences. The channel estimation based upon these known sequences at the users' side provides CSIR which can be used in the detection of DL data.

In LTE TDD frame, even if all data needs to be transmitted in the UL direction, the subframes numbered 0 and 5 in each frame (consisting of 10 subframes) are DL subframes [33] because the BS has to assign resources and provide control information to the users anyway. This means that the users always get the DL channel estimates which we shall be exploiting in our novel CSIT acquisition strategy.

B. Classical CSIT Acquisition in FDD Systems

Pilots transmitted by the BS in the DL direction provide the DL channel information to the users. In an FDD system, UL and DL bandwidths are normally quite far apart and the channel realizations are, in general, independent. Hence the only means to provide the DL channel information to the BS is through explicit feedback of users' DL channel knowledge on the UL resource [12], [10], [14], [15], [16]. For the BS to be able to decode the feedback properly (sent as UL payload), it should first know/estimate the UL channel. To overcome this difficulty, the users first transmit pilot sequences which provide the BS the knowledge of UL channel and then DL channel information is transmitted as UL payload [20].

C. Classical CSIT Acquisition in TDD Systems

If the communication system is operating under TDD mode which is the main focus of this contribution, DL and UL channels are reciprocal. If T_{fb} channel uses are reserved for CSIT acquisition, conventionally a user will transmit pilot sequence of this length, denoted as $\mathbf{x_p} \in \mathbb{C}^{1\times T_{fb}}$, on the UL, as shown in Fig. 1(b) [8] [7]. The signal $\mathbf{Y_p}$ received at the BS is given by

$$\mathbf{Y_p} = \sqrt{P}\,\mathbf{h}\mathbf{x_p} + \mathbf{N_p}, \tag{1}$$

where $\mathbf{N_p} \in \mathbb{C}^{M\times T_{fb}}$ represents the spatio-temporally white Gaussian noise with zero-mean unit-variance entries and $\mathbf{Y_p} \in \mathbb{C}^{M\times T_{fb}}$ is the received signal at M antennas of the BS during this T_{fb}-length training interval. P represents the user's peak power constraint which is equal to the UL signal-to-noise-ratio (SNR) at every BS antenna due to the normalized noise variances. Then simple (UL) channel estimation at the BS will furnish CSIT due to UL and DL channel reciprocity.

III. Optimal Training and Feedback Combining in TDD Systems

The classical training based CSIT acquisition for TDD systems ignores the fact that user knows the DL channel and the CSIT acquisition based only on the quantized feedback for FDD systems cannot use the channel reciprocity (non-existent in FDD systems) whereas in TDD systems both can be exploited at the same time. Working under a constraint of fixed resource available for CSIT acquisition (T_{fb} channel uses and user's power constraint of P), we want to have the best CSIT estimate which fully exploits the reciprocity and the user's channel knowledge simultaneously. We assume perfect channel knowledge at the user's side for ease of exposition[1] and later, imperfect CSIR analysis is carried out in section V-C and simulation results are presented in section VII.

The optimum CSIT acquisition problem in this setting is to inform the receiver (BS) about the fading state known causally at the user. This problem formulation resembles the case in [35] where the authors treated the problem of state information transmission to the receiver for the unfaded case. Let the mean square error in CSIT with respect to (w.r.t.) the true state be selected as the performance metric. Then the input sequence should be designed as a function of the known state $\mathbf{h}$, denoted as $\mathbf{x_i}(\mathbf{h}) \in \mathbb{C}^{1 \times T_{fb}}$, in such a manner as to provide the best estimate (w.r.t the metric) at the BS. Let the received signal at the BS corresponding to the input sequence $\mathbf{x_i}(\mathbf{h})$ be denoted by $\mathbf{Y_i} \in \mathbb{C}^{M \times T_{fb}}$.

$$\mathbf{Y_i} = \sqrt{P}\, \mathbf{h}\mathbf{x_i}(\mathbf{h}) + \mathbf{N_i} \tag{2}$$

$\mathbf{N_i} \in \mathbb{C}^{M \times T_{fb}}$ denotes the spatio-temporally white Gaussian noise at the BS. The BS needs to extract the state information from the observation sequence where it appears not only in the encoded sequence $\mathbf{x_i}(\mathbf{h})$ but also in the state $\mathbf{h}$ of the channel which is used to convey the input sequence $\mathbf{x_i}(\mathbf{h})$. If $\mathcal{O}$ denotes this optimal extraction function, the optimal state detected at the BS is given by

$$\mathbf{h_{opt}} = \mathcal{O}(\mathbf{Y_i}). \tag{3}$$

So the information theoretic optimal CSIT acquisition in this setting requires the optimal encoding of the known state at the transmitter (user) and the optimal state extraction from the encoded information and the state itself at the receiver (BS).

Although the information theoretic problem formulation suggests the fundamental limits of operation, in most of the cases the analysis becomes intractable. And even if the optimal solution is known, it may require large block lengths for coding and infinite delays for quantization and decoding etc, making it impractical to implement such a solution in practical systems. This motivates us to find a practically viable solution which might be suboptimal but still capturing the performance better than the existing solutions. We propose a **novel hybrid two stage CSIT acquisition strategy** which exploits the channel reciprocity and user's channel knowledge at the same time. Working under a constraint of fixed resource available for

[1]In general, the CSIR quality at the users' side is much better. Firstly the DL pilots are global (they are not transmitted per user contrary to the UL pilots) and secondly, the BS can surely pump larger power as compared to small hand-held mobile devices.

CSIT acquisition (T_{fb} channel uses and user's power constraint of P), our strategy consists of dividing this interval in two phases as shown in Fig. 1(c), contrary to the classical UL pilot sequence transmission in Fig. 1(b). The first stage of this hybrid approach, termed as "UL training", is the transmission of normalized training sequence $\mathbf{x_a} \in \mathbb{C}^{1\times T_a}$ from the user to the BS for T_a channel uses and the received signal will be

$$\mathbf{Y_a} = \sqrt{P}\,\mathbf{h}\mathbf{x_a} + \mathbf{N_a}. \tag{4}$$

where $\mathbf{N_a} \in \mathbb{C}^{M\times T_a}$ represents the spatio-temporally white Gaussian noise with zero-mean unit-variance entries and $\mathbf{Y_a} \in \mathbb{C}^{M\times T_a}$ is the received signal at M antennas of the BS during this T_a-length training interval. The optimal training based estimate, denoted as $\hat{\mathbf{h}}_\mathbf{a}$, based upon the observed signal $\mathbf{Y_a}$ and knowing $\mathbf{x_a}$ will be:

$$\hat{\mathbf{h}}_\mathbf{a} = \arg\min_{\mathbf{h}} ||\mathbf{Y_a} - \sqrt{P}\,\mathbf{h}\mathbf{x_a}||^2 \tag{5}$$

The second stage, termed as "quantized feedback", consists of the transmission of quantized channel, already known at the user as a consequence of the DL training. If Q denotes the quantization function, then the quantized version of the channel at the user (the index of the closest codeword in the codebook) is given by $Q(\mathbf{h})$. Afterward user maps this index (sequence of bits) into a sequence of constellation symbols, using the mapping function denoted by S. Let the finite cardinality set of all mapped codewords be denoted by $\mathcal{CB}$. Hence the UL feedback would be

$$\mathbf{x_q} = S(Q(\mathbf{h})), \tag{6}$$

where $\mathbf{x_q} \in \mathbb{C}^{1\times T_q}$ is the T_q dimensional row vector of the normalized constellation symbols. If the user transmits this feedback, the received signal at the BS is given by

$$\mathbf{Y_q} = \sqrt{P}\,\mathbf{h}\mathbf{x_q} + \mathbf{N_q}, \tag{7}$$

where $\mathbf{Y_q}$ and $\mathbf{N_q}$ are $M \times T_q$ matrices of the received signal and the noise respectively at M antennas of the BS during this explicit T_q length feedback interval. This equation reveals the intriguing aspect that the BS needs to acquire $\mathbf{h}$ which appears both as the channel and the transmitted feedback $\mathbf{x_q}$. The BS can try to decode only the quantized channel information based upon the knowledge of $\hat{\mathbf{h}}_\mathbf{a}$ (obtained as in eq. (5) making use of pure training $\mathbf{x_a}$)

$$\hat{\mathbf{h}}_\mathbf{q} = \arg\min_{\mathbf{x_q}\in\mathcal{CB}} ||\mathbf{Y_q} - \sqrt{P}\,\hat{\mathbf{h}}_\mathbf{a}\mathbf{x_q}||^2. \tag{8}$$

The optimal CSIT for this strategy will be obtained by the joint estimation and detection (of $\mathbf{h}$ and $\mathbf{x_q}$ respectively) based upon the observation of $\mathbf{Y_a}$ and $\mathbf{Y_q}$, knowing $\mathbf{x_a}$ and assuming an optimal split between the training and the quantized feedback phases (constrained as $T_a + T_q = T_{fb}$).

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} || \, [\mathbf{Y_a} \; \mathbf{Y_q}] - \sqrt{P}\, \mathbf{h}[\mathbf{x_a} \; S(Q(\mathbf{h}))] \, ||^2 \tag{9}$$

The optimal solution requires a double minimization and does not seem to bear a closed form expression for $\hat{\mathbf{h}}$.

Interestingly even this sub-optimal scheme of splitting the acquisition resource among training and quantization based feedback can be cast in terms of a well-known problem in information theory - the Wyner-Ziv setup [36] which basically treats the problem of decoding with side information at the receiver. Some variants of this problem have been treated where the objective could be some function of message (encoded data) and the side information e.g. [37]. Our proposed hybrid setup will give the problem formulation where the objective is to minimize the distortion in CSIT, decoding the quantized information in the presence of side information from the pilot sequence. As argued earlier, we shall keep our focus on practical signal processing algorithms which might confirm our intuition about CSIT acquisition through this hybrid setup.

IV. Algorithms for Joint Channel Estimation and Feedback Detection

We give three algorithms in this section which separately solve the estimation and the detection problem of the joint minimization of eq. (9). The first two algorithms are iterative which separately solve the estimation and detection problems and iterate till convergence. These algorithms have been closely inspired by [38] which proposes similar algorithms for joint blind estimation and detection for signal separation. We have made modifications for our requirements where data aided channel estimation after the initialization step and the presence of channel as "data" (feedback) make it quite unique. The third algorithm is just the single-shot solution of the joint estimation and detection. Owing to its simplicity, it allows us to further optimize the resource split between training and quantized feedback in the next section. These algorithms assume perfect channel reciprocity as if the calibration were ideal. In case of imperfect reciprocity/calibration, imperfection will add additional noise to the system, the detailed analysis of which goes out of the scope of the present work.

A. Iterative Estimation and Detection

We describe below our algorithm.

Step 1) Initial channel estimation based only upon pilots

$$\hat{\mathbf{h}}_{\mathbf{a}}^{0} = \arg\min_{\mathbf{h}} ||\mathbf{Y_a} - \sqrt{P}\, \mathbf{h}\mathbf{x_a}||^2, \tag{10}$$

which is a simple least squares problem with the solution

$$\hat{\mathbf{h}}_{\mathbf{a}}^{0} = \mathbf{Y_a}\mathbf{x_a}^{\dagger}(\mathbf{x_a}\mathbf{x_a}^{\dagger})^{-1}\frac{1}{\sqrt{P}}. \tag{11}$$

$$i = 1 \tag{12}$$

Superscript denotes the iteration number.

Step 2) At iteration i, do enumeration over all the codes in the codebook assuming that the channel $\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}$ is perfectly known.

$$\hat{\mathbf{x}}_{\mathbf{q}}^{i} = \arg\min_{\mathbf{x_q} \in \mathcal{CB}} ||\mathbf{Y_q} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\mathbf{x_q}||^2 \tag{13}$$

Step 3) Regenerate extended pilot sequence $\mathbf{x_{ext}}$ (pilots and detected feedback)

$$\mathbf{x_{ext}}^{i} = [\mathbf{x_a}\ \hat{\mathbf{x}}_{\mathbf{q}}^{i}]. \qquad \mathbf{Y_{ext}} = [\mathbf{Y_a}\ \mathbf{Y_q}]. \tag{14}$$

Step 4) Channel estimation based upon extended pilots (i.e. knowing $\mathbf{x_{ext}}^{i}$)

$$\hat{\mathbf{h}}_{\mathbf{a}}^{i} = \arg\min_{\mathbf{h}} ||\mathbf{Y_{ext}} - \sqrt{P}\,\mathbf{h}\mathbf{x_{ext}}^{i}||^2 \tag{15}$$

$$\hat{\mathbf{h}}_{\mathbf{a}}^{i} = \mathbf{Y_{ext}}\mathbf{x_{ext}}^{\dagger i}(\mathbf{x_{ext}}^{i}\mathbf{x_{ext}}^{\dagger i})^{-1}\frac{1}{\sqrt{P}} \tag{16}$$

Step 5) If $\hat{\mathbf{x}}_{\mathbf{q}}^{i} \neq \hat{\mathbf{x}}_{\mathbf{q}}^{i-1}$ or $\hat{\mathbf{h}}_{\mathbf{a}}^{i} \neq \hat{\mathbf{h}}_{\mathbf{a}}^{i-1}$, $i = i + 1$ and go to Step 2.

The final channel estimate $\hat{\mathbf{h}}$ is the channel vector corresponding to $\hat{\mathbf{x}}_{\mathbf{q}}^{i}$ in the codebook.

Theorem 1 (Convergence for Iterative Estimation and Detection Algorithm): Let $\hat{\mathbf{h}}_{\mathbf{a}}^{i}$ be the estimated channel and $\hat{\mathbf{x}}_{\mathbf{q}}^{i}$ be the detected feedback, both at i-th iteration of the iterative estimation and detection algorithm. Let the residual function $f\left(\hat{\mathbf{h}}_{\mathbf{a}}, \mathbf{x_{ext}}; \mathbf{Y_{ext}}\right) \triangleq ||\mathbf{Y_{ext}} - \sqrt{P}\ \hat{\mathbf{h}}_{\mathbf{a}}\mathbf{x_{ext}}||^2$ be selected as the descent function for this algorithm. Then there exists some j such that for any $i \geq j$, $\hat{\mathbf{x}}_{\mathbf{q}}^{i} = \hat{\mathbf{x}}_{\mathbf{q}}^{j}$ and $\hat{\mathbf{h}}_{\mathbf{a}}^{i} = \hat{\mathbf{h}}_{\mathbf{a}}^{j}$.

Proof: The residual descent function $f\left(\hat{\mathbf{h}}_{\mathbf{a}}, \mathbf{x}_{\mathbf{ext}}; \mathbf{Y}_{\mathbf{ext}}\right) = ||\mathbf{Y}_{\mathbf{ext}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}\mathbf{x}_{\mathbf{ext}}||^2$ is clearly non-negative and continuous. Considering the residual function at i-th iteration:

$$\begin{aligned}
f\left(\hat{\mathbf{h}}_{\mathbf{a}}^{i}, {\mathbf{x}_{\mathbf{ext}}}^{i}; \mathbf{Y}_{\mathbf{ext}}\right) &\overset{a}{=} ||\mathbf{Y}_{\mathbf{ext}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i}{\mathbf{x}_{\mathbf{ext}}}^{i}||^2 \\
&\overset{b}{=} \min_{\mathbf{h}} ||\mathbf{Y}_{\mathbf{ext}} - \sqrt{P}\,\mathbf{h}{\mathbf{x}_{\mathbf{ext}}}^{i}||^2 \\
&\overset{c}{\leq} ||\mathbf{Y}_{\mathbf{ext}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}{\mathbf{x}_{\mathbf{ext}}}^{i}||^2 \\
&\overset{d}{=} ||\mathbf{Y}_{\mathbf{a}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\mathbf{x}_{\mathbf{a}}||^2 + ||\mathbf{Y}_{\mathbf{q}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\hat{\mathbf{x}}_{\mathbf{q}}^{i}||^2 \\
&\overset{e}{=} ||\mathbf{Y}_{\mathbf{a}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\mathbf{x}_{\mathbf{a}}||^2 + \min_{\mathbf{x}_{\mathbf{q}} \in \mathcal{CB}} ||\mathbf{Y}_{\mathbf{q}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\mathbf{x}_{\mathbf{q}}||^2 \\
&\overset{f}{\leq} ||\mathbf{Y}_{\mathbf{a}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\mathbf{x}_{\mathbf{a}}||^2 + ||\mathbf{Y}_{\mathbf{q}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}\hat{\mathbf{x}}_{\mathbf{q}}^{i-1}||^2 \\
&\overset{g}{=} ||\mathbf{Y}_{\mathbf{ext}} - \sqrt{P}\,\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}{\mathbf{x}_{\mathbf{ext}}}^{i-1}||^2 \\
&\overset{h}{=} f\left(\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}, {\mathbf{x}_{\mathbf{ext}}}^{i-1}; \mathbf{Y}_{\mathbf{ext}}\right) \qquad (17)
\end{aligned}$$

Equalities d and g make use of the property of the Frobenius norm [39]. The set of equations above shows that each single iteration of the algorithm over estimation and detection causes to monotonically reduce the residual function unless iterates converge. This monotonic reduction of the descent function, its non-negativity and the fact that $\mathbf{x}_{\mathbf{q}}$ belongs to a finite set (codes of the codebook) and hence corresponding iterates of the estimation subproblem are also finite prove the convergence of this algorithm to the locally optimal solution in a finite number of steps. The globally optimal solution is achieved by having a good initial point which depends upon the training part as confirmed by our simulations. ■

B. *Simplified Iterative Estimation and Detection*

This algorithm is very similar to the previous algorithm in essence but the difference arises at the detection step. The second step of the previous algorithm, the ML detection of the quantized code from the codebook, is computationally quite onerous, especially for codebooks with large cardinality. So we replace this enumeration step with least squares detection followed by mapping on the codebook. So the Step 2 of the previous algorithm gets replaced by two sub-steps.

Step 2-A) At iteration i, do LS detection of the quantized feedback assuming $\hat{\mathbf{h}}_{\mathbf{a}}^{i-1}$ as the perfectly known channel

$$\hat{\mathbf{x}}_{\mathrm{LS}}^{i} = (\hat{\mathbf{h}}_{\mathbf{a}}^{\dagger i-1}\hat{\mathbf{h}}_{\mathbf{a}}^{i-1})^{-1}\hat{\mathbf{h}}_{\mathbf{a}}^{\dagger i-1}\mathbf{Y}_{\mathbf{q}}\frac{1}{\sqrt{P}}. \qquad (18)$$

Step 2-B) Do hard detection of LS estimate $\hat{\mathbf{x}}^i_{\mathrm{LS}}$ on the constellation symbols which will map the LS detected channel feedback to the nearest code in the codebook.

$$\hat{\mathbf{x}}^i_{\mathbf{q}} = \mathrm{HardDetection}(\hat{\mathbf{x}}^i_{\mathrm{LS}}) \tag{19}$$

LS followed by hard detection significantly reduces the computational complexity of this algorithm but this step of hard detection prevents the analytical convergence proof.

C. Single-Shot Estimation and Detection

This is the simplest and the fastest algorithm for the joint estimation and detection problem where the channel estimation and the feedback detection are performed (separately) only once.

Step 1) Channel estimation based only upon the pilots

$$\hat{\mathbf{h}}_{\mathbf{a}} = \arg\min_{\mathbf{h}} ||\mathbf{Y}_{\mathbf{a}} - \sqrt{P}\,\mathbf{h}\mathbf{x}_{\mathbf{a}}||^2. \tag{20}$$

One can employ the MMSE criteria instead of LS.

Step 2) Detection of the feedback $\mathbf{x}_{\mathbf{q}}$ assuming channel $\hat{\mathbf{h}}_{\mathbf{a}}$ is perfectly known. This detection problem can be solved either by enumerating all the codewords as in the first algorithm or by simple LS as in the second algorithm or even by applying MMSE filter.

V. Outage Based Training and Feedback Optimization

A. Definitions and Initial Setup

The solution for the optimal CSIT estimate, $\hat{\mathbf{h}}$ in eq. (9), requires joint estimation and detection. Furthermore, the fixed resource (T_{fb} channel uses) needs to be optimally split between training and feedback. Even if, as a simplification, we focus separately on training based estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ (given in eq. (5)) and digital feedback based estimate $\hat{\mathbf{h}}_{\mathbf{q}}$ (given in eq. (8)), three questions still need to be answered:

1) how the fixed CSIT acquisition interval T_{fb} should be split between training and feedback?
2) what should be the rate of the quantized feedback?
3) how the two estimates should be combined to get the final estimate?

We use the minimization of the mean-square error (MSE) of the final CSIT (defined below) as the criterion to determine the resource split and the rate of the quantized feedback, thus answering the first two questions for which we give the proper framework in the next subsection. This choice is justified as we detailed in section II that the CSIT reliability directly translates into system wide throughput [10]. Furthermore, we propose to use the quantized feedback based estimate $\hat{\mathbf{h}}_{\mathbf{q}}$ as the final CSIT estimate $\hat{\mathbf{h}}$

due to the better MSE decay behavior associated to the quantized feedback as an answer to the third question. We revisit this statement in section V-D. It may appear that the training based estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ is not used properly but in reality quantized feedback $\mathbf{x}_{\mathbf{q}}$, which provides $\hat{\mathbf{h}}_{\mathbf{q}}$, is decoded based upon this training based estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ hence combining is implicitly achieved.

The optimization framework is based upon the single shot estimation and detection algorithm, proposed in section IV-C. It consists of first providing a training based estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ to the BS in the training interval of T_a channel uses. In the second interval of T_q channel uses, the user sends the quantized version of its unit-norm channel direction information (CDI) vector which we assume to be perfectly known at the user. The size of the codebook employed will depend upon the rate of the feedback. It was pointed out in [40], [41] and [30] that when channel fades cannot be averaged out either because of stringent decoding delay constraints or because of a very slowly fading channel, the classical notion of Shannon capacity bears no meaning and there is an error probability associated to each rate. This rate is called the outage rate and the associated probability is the probability that the channel can not support this rate. For this CSIT quantized feedback transmission over a single channel realization, a case with the most stringent decoding delay scenario (or equivalently a very slow fading channel where transmission sees only one realization), deep channel fades (causing outage) are the typical error events [30]. Following [42], we ignore feedback decoding errors beyond a system outage event.

We define the "outage" as an event when the channel realization and the quality of the training based estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ (a function of T_a) do not allow the BS to successfully decode the feedback information. Let $\epsilon(T_a, b)$ be the outage probability when quantized feedback is transmitted at a rate of b bits per channel use. Thus b is the $\epsilon(T_a, b)$-outage rate of the UL channel [30]. So the user can send a total of $B = bT_q$ feedback bits at $\epsilon(T_a, b)$ outage. Although the constellations used in practice have 2^b points where b must be a positive integer, for the time being we relax this restriction and allow positive real values for b.

We define the squared CDI error as the sine squared of the angle (θ) between the true channel direction vector $\bar{\mathbf{h}}$ and the BS estimated direction vector $\bar{\hat{\mathbf{h}}}$, denoted as $\sigma^2(\mathbf{h}, \hat{\mathbf{h}})$.

$$\sigma^2(\mathbf{h}, \hat{\mathbf{h}}) \triangleq \sin^2(\theta) = 1 - \cos^2(\theta) = 1 - |\bar{\mathbf{h}}^\dagger \bar{\hat{\mathbf{h}}}|^2 \tag{21}$$

Further the MSE of CSIT is defined to be the expected value of the squared CDI error at the transmitter and denoted as σ^2. Although it is a slight abuse of notation but it has been shown that the CDI plays a vital role both for single-user and multi-user scenarios [10], [11] and secondly CDI makes the major feedback burden in CSIT compared to channel norm scalar.

For the quantization of M-dimensional unit-norm CDI at the user, we employ random vector quantization (RVQ). For RVQ, the exact expression for the mean-square quantization error σ_q^2 has been given in [43], [10] as

$$\sigma_q^2 = 2^B \beta\left(2^B, \frac{M}{M-1}\right), \tag{22}$$

where B is the total number of feedback bits (i.e. the codebook consists of 2^B codes) and β represents the beta function which is defined in terms of the Gamma function as $\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. However it turns out that a simple and tight upper bound given in reference [10] suffices:

$$\sigma_q^2 \leq 2^{\frac{-B}{M-1}}. \tag{23}$$

B. Optimal Resource Split between Training and Quantized Feedback

Theorem 2 (The minimization of the MSE of CSIT): Under the training and feedback combining strategy, the MSE of CSIT σ^2 is minimized as a result of the following optimization governing the fixed resource (T_{fb}) split between the training T_a and the quantized feedback interval T_q and the outage rate b:

$$\sigma^{2*} = \underset{T_a, b}{min} \left[2^{\frac{-b(T_{fb}-T_a)}{M-1}} + \epsilon(T_a, b)\right] \tag{24}$$

The constraints for this minimization are:

$$1 \leq T_a \leq T_{fb} \quad \text{and} \quad 0 \leq b \tag{25}$$

The outage probability in the feedback interval $\epsilon(T_a, b)$ and the feedback rate b are linked by the relation:

$$b = \log\left(1 + \frac{P^2 T_a}{2(P + PT_a + 1)} F^{-1}(\epsilon(T_a, b))\right), \tag{26}$$

where P is the user's power constraint and $F^{-1}(.)$ is the inverse of the standard cumulative distribution function (CDF) of χ^2_{2M} distributed variable.

Proof: The proof consists of two parts. First we show the argument of minimization to be an upper bound on the MSE of CSIT and in the second part, the relation between $\epsilon(T_a, b)$ and b is derived.

Upper bound on the MSE of CSIT: During the feedback phase, when the channel is not in outage and the BS is able to decode the feedback correctly, there is only quantization error in the final CSIT estimate. On the other hand, when the channel is in outage (happens with probability $\epsilon(T_a, b)$), the BS cannot decode the feedback information. Hence the MSE of CSIT σ^2 can be written as

$$\begin{aligned} \sigma^2 &= (1 - \epsilon(T_a, b))\, \sigma_q^2 + \epsilon(T_a, b)\, \mathbb{E}\sigma^2_{\bar{\mathbf{h}} \neq \bar{\hat{\mathbf{h}}}}(\mathbf{h}, \hat{\mathbf{h}}) \\ &\leq (1 - \epsilon(T_a, b))\, \sigma_q^2 + \epsilon(T_a, b) \\ &\leq \sigma_q^2 + \epsilon(T_a, b), \end{aligned} \tag{27}$$

where σ_q^2 is the mean-square quantization error and $\sigma^2_{\bar{\mathbf{h}}\neq\bar{\hat{\mathbf{h}}}}(\mathbf{h}, \hat{\mathbf{h}})$ represents the MSE of CSIT when the channel is in outage (which means a feedback error occurs). The first inequality is obtained as $\mathbb{E}\sigma^2_{\bar{\mathbf{h}}\neq\bar{\hat{\mathbf{h}}}}(\mathbf{h}, \hat{\mathbf{h}})$ is upper-bounded by 1. Putting the value of σ_q^2 from eq. (23) using $B = bT_q$ and $T_{fb} = T_a + T_q$ in eq. (27), we get the desired upper bound of the MSE of CSIT as

$$\sigma^2 \leq 2^{\frac{-b(T_{fb}-T_a)}{M-1}} + \epsilon(T_a, b), \tag{28}$$

which concludes the first part of our proof.

Significance of the MSE bound: The MSE bound of the CSIT eq. (28) is the desired performance metric. Its minimization gives us the optimal values for T_a, T_q and b (the number of feedback bits per channel use - this parameter governs the constellation size and hence the quantization error) for a fixed resource T_{fb}. This bound shows us the basic trade-off involved. If the total number of feedback bits $B = bT_q$ is made large (either by choosing a large rate b per channel use in the feedback channel or by making T_q large), it will allow the user to select a larger codebook (with 2^B codewords) and hence the quantization error will be negligible. But this strategy will plague the final CSIT estimation error by introducing a lot of outage events (due to large b or poor channel estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ caused by small $T_a = T_{fb} - T_q$). On the other hand for a small number of total feedback bits B, the degradation due to outage probability will fade away, but there will be fewer codewords in the codebook and hence a large quantization error.

The relation of b and $\epsilon(T_a, b)$: Pilot sequence transmission from the user to the BS for an interval of length T_a, given in eq. (4), can be equivalently written in a simplified form as

$$\mathbf{y_a} = \sqrt{PT_a}\,\mathbf{h} + \mathbf{n_a}, \tag{29}$$

where P is the user's power constraint and $\mathbf{y_a}, \mathbf{h}, \mathbf{n_a}$ are the received signal, the channel vector and the noise respectively, all column vectors of dimension M. The BS can make MMSE estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ of the channel $\mathbf{h}$ as

$$\hat{\mathbf{h}}_{\mathbf{a}} = \frac{\sqrt{PT_a}}{PT_a + 1}\mathbf{y_a}. \tag{30}$$

As the i.i.d. channel entries are standard Gaussian, the MMSE estimation error $\tilde{\mathbf{h}}_{\mathbf{a}} = \mathbf{h} - \hat{\mathbf{h}}_{\mathbf{a}}$ has also Gaussian i.i.d. entries as $\tilde{\mathbf{h}}_{\mathbf{a}} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_a^2 \mathbf{I_M}\right)$ and the MSE per channel coefficient σ_a^2 is given by

$$\sigma_a^2 = \frac{1}{PT_a + 1}. \tag{31}$$

Similarly the estimate $\hat{\mathbf{h}}_{\mathbf{a}}$ has Gaussian i.i.d. entries and is distributed as $\hat{\mathbf{h}}_{\mathbf{a}} \sim \mathcal{CN}\left(\mathbf{0}, \frac{PT_a}{PT_a+1}\mathbf{I_M}\right)$.

Now we focus our attention on the quantized feedback interval of the CSIT acquisition, given in eq. (7). The signal received during one symbol interval of this phase is given by

$$\mathbf{y_q} = \sqrt{P}\,\mathbf{h}x_q + \mathbf{n_q}, \tag{32}$$

where x_q represents the scalar feedback symbol transmitted by the user and $\mathbf{y_q}, \mathbf{h}, \mathbf{n_q}$ are M-dimensional column vectors representing respectively the observed signal, the channel and the noise for this particular symbol interval. To decode this information, the BS uses the estimate $\hat{\mathbf{h}}_\mathbf{a}$ that it developed during the training phase. So the above equation can be written as

$$\mathbf{y_q} = \sqrt{P}\,\hat{\mathbf{h}}_\mathbf{a}x_q + \sqrt{P}\,\tilde{\mathbf{h}}_\mathbf{a}x_q + \mathbf{n_q}. \tag{33}$$

The average effective signal-to-noise-ratio (denoted as $\mathrm{SNR}_{\mathrm{eff}}$) at the BS during the feedback interval relegating the imperfect channel estimate portion of the signal into noise and treating $\hat{\mathbf{h}}_\mathbf{a}$ as the perfectly known channel is given by:

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{P||\hat{\mathbf{h}}_\mathbf{a}||^2}{P\sigma_a^2 + 1}. \tag{34}$$

Plugging in the value of σ_a^2 from eq. (31), $\mathrm{SNR}_{\mathrm{eff}}$ will become

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{P||\hat{\mathbf{h}}_\mathbf{a}||^2}{\frac{P}{PT_a+1} + 1}. \tag{35}$$

We can do a small change of variable as $\frac{2(PT_a+1)}{PT_a}||\hat{\mathbf{h}}_\mathbf{a}||^2$ represents a standard chi-square random variable having $2M$ degrees of freedom (DOF), denoted as χ^2_{2M}. So the $\mathrm{SNR}_{\mathrm{eff}}$ becomes

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{P^2T_a}{2(P+PT_a+1)}\chi^2_{2M}. \tag{36}$$

The outage probability $\epsilon(T_a, b)$ during this feedback interval corresponding to the outage rate b bits per channel use can be written as

$$\begin{aligned} \epsilon(T_a, b) &= \mathbb{P}\left[\log\left(1 + \mathrm{SNR}_{\mathrm{eff}}\right) \le b\right] \\ &= \mathbb{P}\left[\log\left(1 + \frac{P^2T_a}{2(P+PT_a+1)}\chi^2_{2M}\right) \le b\right], \end{aligned} \tag{37}$$

where $\mathbb{P}$ denotes the probability of an event. This relation can be inverted to obtain the outage rate b corresponding to the outage probability $\epsilon(T_a, b)$, as given below

$$b = \log\left(1 + \frac{P^2T_a}{2(P+PT_a+1)}F^{-1}(\epsilon(T_a, b))\right), \tag{38}$$

where $F^{-1}(.)$ is the inverse of the CDF of χ^2_{2M} distributed variable. This concludes the proof. ∎

The analytical solution to the minimization in Theorem 2 does not bear closed form expression but its numerical optimization is quite trivial.

C. Resource Split with Imperfect CSIR

Let the MSE in the CDI at the user be σ_r^2 and the MS quantization error be denoted by σ_q^2 as earlier. The relation of outage probability and the outage rate in the feedback interval still holds but the expression for the MSE of CSIT changes. We make use of eq. (27) from the previous subsection which gives a bound on the MSE of CSIT splitting two cases when feedback is received correctly or not. With correct feedback detection, now the MSE of CSIT will carry the impact of quantization and imperfect CSIR. Denoting this MSE of quantized imperfect CSIR by σ_Q^2, eq. (27) becomes

$$\sigma^2 \leq \sigma_Q^2 + \epsilon(T_a, b). \tag{39}$$

Now we need to specify σ_Q^2 in terms of quantization error and CSIR imperfection. Let $\mathbf{h_r}$ denote the imperfect CSIR for true channel $\mathbf{h}$ and $\mathbf{h_q}$ be the unit-norm quantized channel at the user's side obtained by quantization of $\mathbf{h_r}$. These vectors and the angles they subtend with each other have been plotted in Fig. 2. Let the angle between $\mathbf{h_q}$ and the true CDI ($\mathbf{h}$) be denoted by θ_Q, then the MSE in the direction

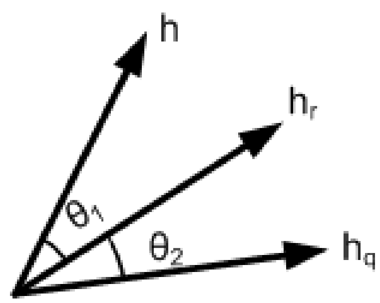


Fig. 2. True channel ($\mathbf{h}$), CSIR ($\mathbf{h_r}$) and user's quantized CDI ($\mathbf{h_q}$).

of true channel and its quantized version is given by

$$\begin{aligned}
\sigma_Q^2 &\stackrel{a}{=} \mathbb{E}\sigma^2(\mathbf{h}, \mathbf{h_q}) = \mathbb{E}\sin^2(\theta_Q) \\
&\stackrel{b}{\leq} \mathbb{E}\sin^2(\theta_1 + \theta_2) \\
&\stackrel{c}{=} \mathbb{E}\left(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2\right)^2 \\
&\stackrel{d}{=} \mathbb{E}\left(\sin^2\theta_1\cos^2\theta_2 + \cos^2\theta_1\sin^2\theta_2 + 2\sin\theta_1\sin\theta_2\cos\theta_1\cos\theta_2\right) \\
&\stackrel{e}{\leq} \mathbb{E}\left(\sin^2\theta_1 + \sin^2\theta_2 + 2\sin\theta_1\sin\theta_2\right) \\
&\stackrel{f}{=} \mathbb{E}\sin^2\theta_1 + \mathbb{E}\sin^2\theta_2 + 2\mathbb{E}\sin\theta_1\mathbb{E}\sin\theta_2 \\
&\stackrel{g}{\leq} \sigma_r^2 + \sigma_q^2 + 2\sqrt{\mathbb{E}\sin^2\theta_1}\sqrt{\mathbb{E}\sin^2\theta_2} \\
&\stackrel{h}{=} \sigma_r^2 + \sigma_q^2 + 2\sqrt{\sigma_r^2}\sqrt{\sigma_q^2} \\
&\stackrel{i}{=} (\sigma_r + \sigma_q)^2
\end{aligned} \tag{40}$$

Inequality b follows as $\theta_Q \leq \theta_1 + \theta_2$ and e follows as $\cos\alpha \leq 1$ for $\alpha \leq \pi/2$. f uses the independence between the angles θ_1 and θ_2 whereas inequality g uses the Jensen's inequality. Plugging this bound of σ_Q^2 in eq. (39), we get a new upper bound of the MSE of CSIT for the case of imperfect CSIR.

$$\sigma^2 \leq (\sigma_r + \sigma_q)^2 + \epsilon(T_a, b) \tag{41}$$

This MSE bound needs to be optimized to determine the optimal resource split and the feedback rate in the case of imperfect CSIR.

D. MSE Decay of Training and Quantized feedback based CSIT

In this subsection, we provide some intuition of why quantized feedback is expected to perform better than simple pilot transmission. Eq. (31) $\sigma_a^2 = \frac{1}{PT_a+1}$ shows that the MSE of the channel estimate obtained through UL training decays linearly with the length of the training interval .

To see the decay rate of the quantized feedback with respect to the feedback interval, we reproduce Eq. (28)

$$\sigma^2 \leq 2^{\frac{-bT_q}{M-1}} + \epsilon(T_a, b). \tag{42}$$

This gives the MSE of CSIT through quantized digital feedback. Suppose for a moment that the outage is negligible (actually the optimization in theorem 2 yields the outage and the quantization error approximately equal) and there is only quantization error where the size of the codebook used for quantization mainly depends upon b, the bits that can be sent in the UL direction. With some sacrifice of rigour and abuse of notation, if we assume that the ergodic rate can be transmitted in the UL direction as if the channel is noise only channel, i.e., $b = \log(1+P)$, the MSE becomes

$$\sigma^2 \leq 2^{\frac{-T_q \log(1+P)}{M-1}} = \frac{1}{(1+P)^{\frac{T_q}{M-1}}}. \quad (43)$$

This equation shows the exponential decay of the MSE of quantized CSIT w.r.t. the length of the feedback interval as compared to the linear decay obtained in the MSE of training based CSIT. This makes the quantized feedback based approach to perform better than simple training based scheme as confirmed by simulations in section VII. If minimal resources are available for transmission, it was established that the simple analog transmission (pilot transmission) performs as good as the digital transmission [34].

VI. Optimization Setup With Practical Constellations

In the previous optimization procedure, we had relaxed the restriction of practical constellations and allowed any positive real value for the feedback rate b bits per channel use. In practical communication systems, the constellations used always have number of points equal to an integer power of 2, i.e., b can only take an integer value. We propose two simple strategies in the following subsections to handle this issue which arises due to this limitation of practical constellations.

A. Resource Split Optimization for a Fixed Constellation

We can optimize the MSE of CSIT for a fixed constellation, i.e. for a fixed feedback rate b. In this case, the outage rate based optimization setup, built in the previous section, remains operational except that b is no more an optimization variable but a fixed parameter corresponding to the chosen constellation. Thus b will assume the values of 2 and 4 for QPSK and 16-QAM, respectively, although any other constellation can be selected. The minimization of the MSE of CSIT will give the optimal resource split tailored for the particular constellation. Hence the perfect CSIR objective function for a fixed constellation (fixed value of b) becomes:

$$\underset{T_a}{min} \left[2^{\frac{-b(T_{fb}-T_a)}{M-1}} + \epsilon(T_a, b) \right] \quad (44)$$

where $T_{fb} = T_a + T_q$ and b are fixed, and b and $\epsilon(T_a, b)$ are related as in Theorem 2. This minimization gives us the optimal value of training length T_a which should be used to get the minimum MSE of CSIT for this particular constellation (fixed b) under fixed values of M, P and T_{fb}. This restriction of fixed constellation brings in some limitations. For example, the use of smaller constellation like QPSK at very high SNR will not be beneficial as CSIT error will stay bounded due to the fixed cardinality of the codebook (hence quantization error will be non-diminishing as a function of SNR) even for asymptotically large values of SNR.

B. Using Real Values of b with Extra Parity Bits

The other way to resolve the issue of discrete practical constellations is through the use of channel coding. This allows us to use positive real values for b, obtained from the original optimization setup. The only restriction, we impose, is that B should take an integer value which can be obtained by using ceiling or floor operation on the product bT_q. Now this B governs the cardinality of the codebook. The actual constellation, which is used to send feedback, is the one larger than that dictated by b, among the available constellations. Let the rate of that constellation be denoted by b_c. Hence the number of total bits, which will be sent in the feedback phase, is $B_c = b_c T_q$ where $B_c > B$ as $b_c > b$. All the extra bits $B_c - B$ in the feedback phase are used as parity bits. So one can employ either linear block codes or convolutional codes with an appropriate rate so as to convert B information (true channel feedback) bits into B_c coded bits. One advantage of using convolutional codes is that puncturing can give more flexibility for rate matching. Now these B_c bits are sent in the digital feedback phase. As the outage rate b is less than the rate b_c of the constellation chosen, the use of larger constellation will give rise to increase in the number of erroneous coded bits. The number of errors will grow large in direct proportion to the difference $B_c - B$. On the other hand, all the extra feedback bits $B_c - B$ are the parity bits and when decoding will be performed at the BS, the capability of this coding/decoding operation to combat the channel errors (introduced in the quantized feedback) is also proportional to this difference, hence compensating the negative impact of using larger constellation.

VII. Simulation Results

Our simulation environment consists of a BS with $M = 4$ antennas and a single user with a single antenna. The channel model and the frame structure are the same as described in Section II. The feedback interval T_{fb} is fixed to 20 channel uses for all simulations. The results with hybrid combining of training and quantized feedback use iterative algorithms proposed in section IV for joint estimation and detection whereas the resource split is computed using theorem 2.

A. Optimization Results for Continuous Constellations

First we present the results when the feedback rate b is not constrained to be an integer and can assume any positive real value. The optimization of the objective function, given in section V, gives us the values for the optimal training length T_a and the optimal feedback rate b for various values of user's power constraint, which is equal to the UL SNR as the noise at every BS antenna has been normalized to have unit variance. Knowing the values of $\epsilon(T_a, b)$ and T_q, computed based upon the optimal values of T_a and

b, allows us to compute the upper bound of the final CSIT error eq. (28). These values have been plotted in dB scale in Fig. 3. For comparison purpose, we have also plotted the MSE of CSIT with classical

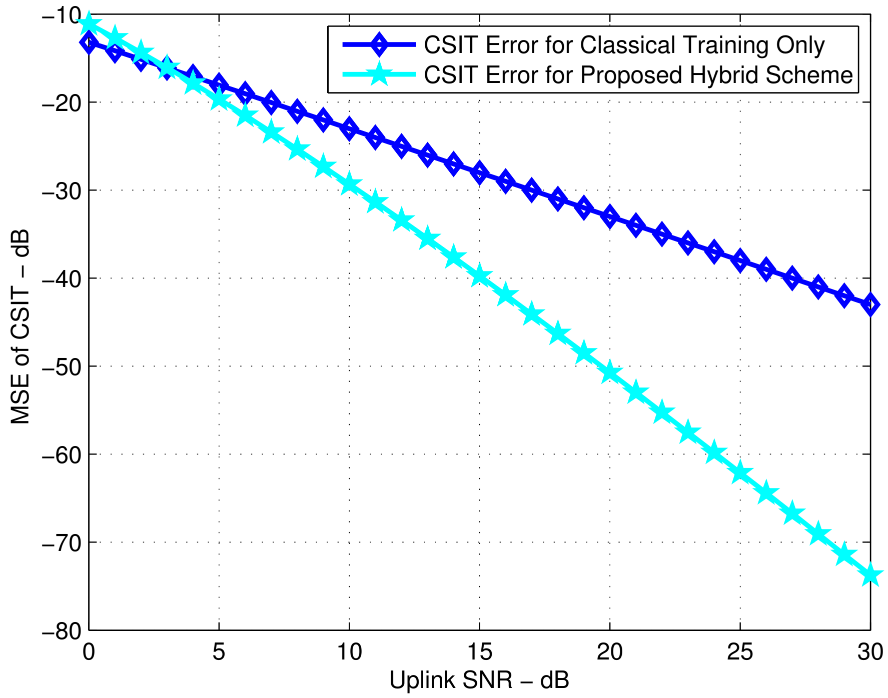


Fig. 3. Mean-Square CSIT Errors with perfect CSIR: $T_{fb} = 20$ and $M = 4$. The novel hybrid scheme performs much better than the classical training based CSIT acquisition. Gains are significant even with naive use of practical constellations without any coding.

training based estimation. This plot clearly shows the interest for our hybrid two-staged CSIT acquisition strategy as, from medium to large SNR values, CSIT error incurred by this scheme is much less than the error obtained by training based only CSIT acquisition. Only at very low SNR values, this two stage scheme performs worse than the classical training scheme.[2] This happens as we have restricted the final estimate to come from the digital feedback. Here the total feedback resource (SNR and T_{fb}) does not allow transmission of sufficient number of bits through the channel so quantization error is quite large.

[2]The proposed CSIT acquisition strategy is suitable from moderate to large values of SNRs. This work focuses on CSIT acquisition for DL multi-user multiplexing. This regime has shown promising gains kicking in from medium to large SNR values. For very low values of SNR, normally diversity approaches will be preferred over user multiplexing where the presence of CSIT is beneficial but not the prime requirement.

This gets aggravated due to the poor training based estimate based upon which these bits are decoded, further degrading the performance. This degradation can be easily avoided by selecting an SNR threshold below which traditional training based scheme should be employed.

To see the optimal split between training and quantized feedback, we have plotted the optimal values of training length T_a, corresponding values of quantized feedback interval T_q and the optimal feedback rate b in Fig. 4.

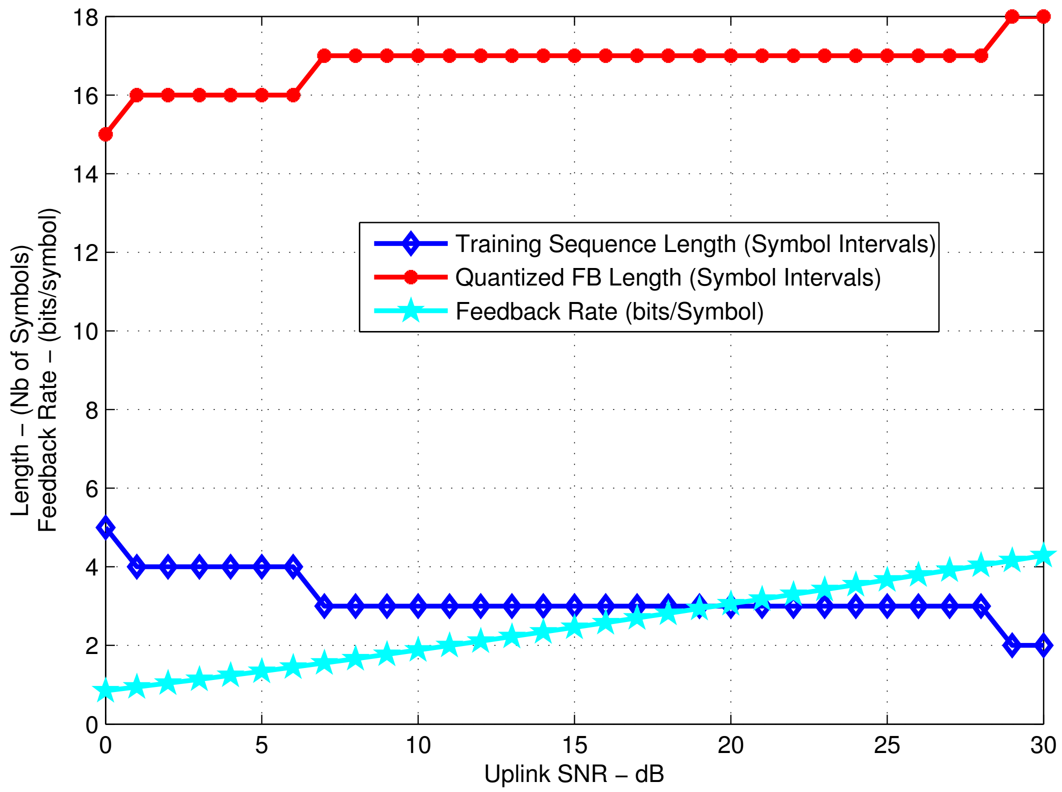


Fig. 4. Optimal Lengths and Outage Rate: $T_{fb} = 20$ and $M = 4$. With increase in SNR, both the length of the quantized feedback interval T_q and the outage rate b increase gradually.

We have plotted the results of the MSE of CSIT obtained through hybrid combining for the case of imperfect CSIR in Fig. 5. The results have been plotted for various levels of CSIR quality. The proposed hybrid scheme shows much better performance than the classical scheme for reasonable quality CSIR. We have plotted the curves when CSIR is 10, 20 and 30 dB better than the quality of CSIT obtained using classical scheme. Finite CSIR quality might lead to the saturation of CSIT at high UL SNRs. Quite logically CSIT acquisition resources should be invested on quantized feedback only when its quality is better than the pilot based only CSIT.

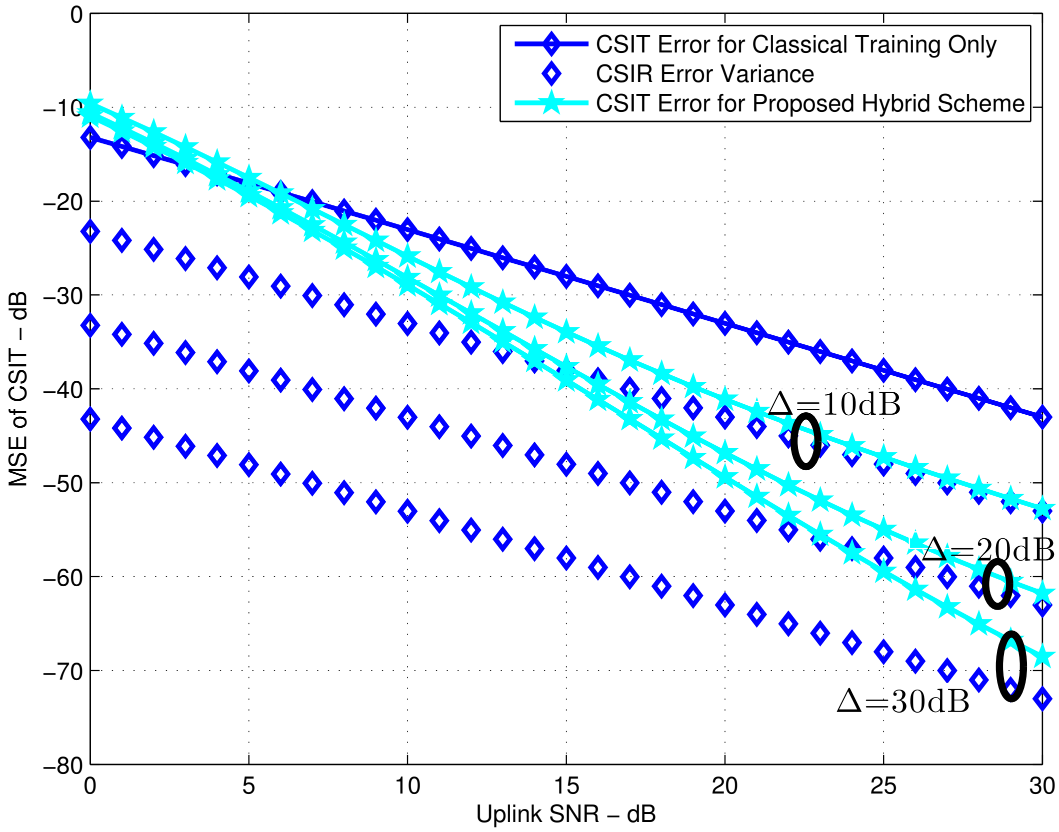
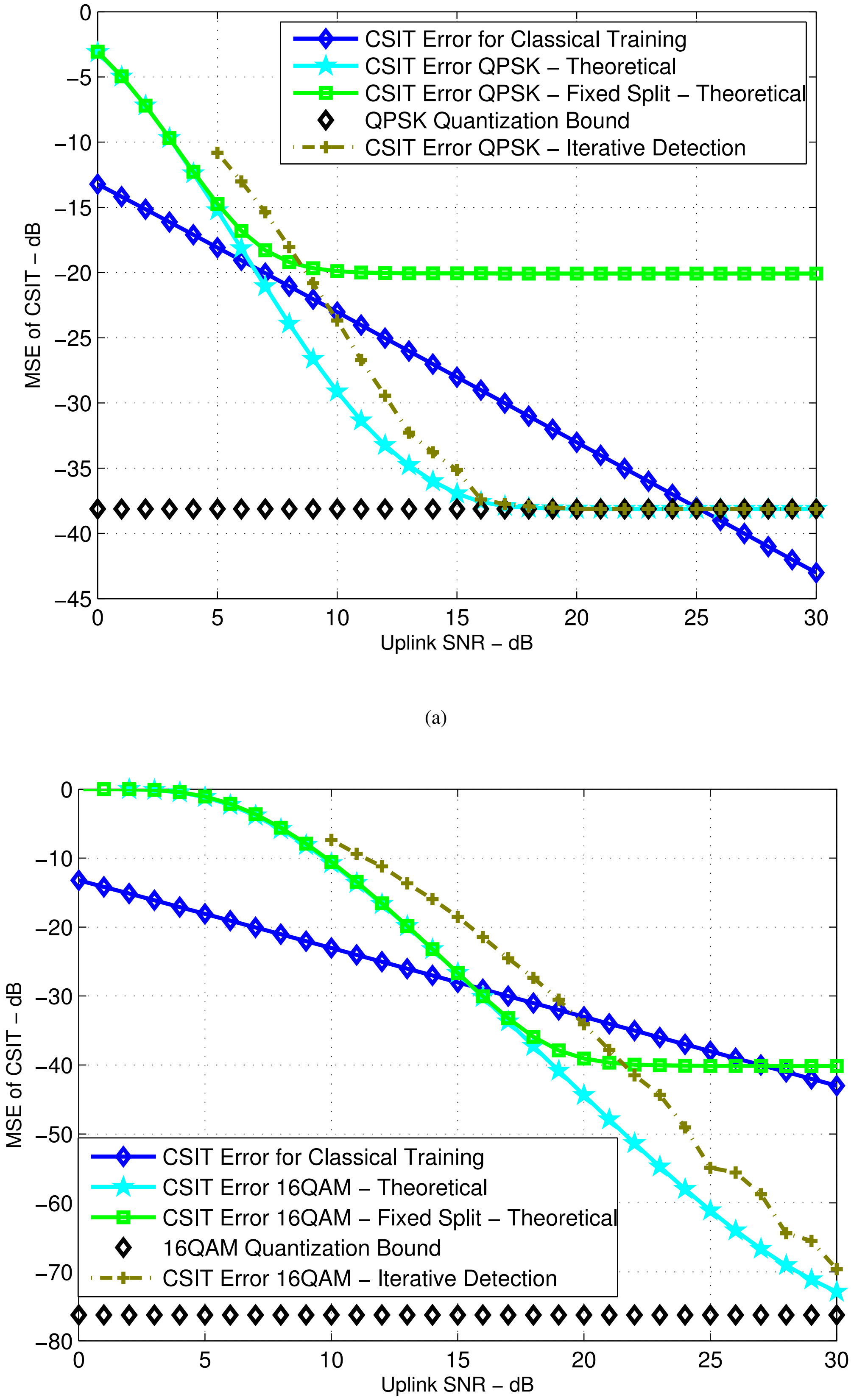


Fig. 5. Mean-Square CSIT Errors with imperfect CSIR: $T_{fb} = 20$ and $M = 4$. The novel hybrid scheme outperforms the classical scheme even with imperfect CSIR. Δ denotes the difference of the quality of CSIR and CSIT.

B. Optimization Results for Discrete Constellations with Perfect CSIR

In this section, we present simulation results when fixed constellations QPSK and 16-QAM are used for quantized feedback transmission. Here the feedback rate b becomes fixed corresponding to the fixed constellation (2 for QPSK and 4 for 16-QAM) and the optimization is carried only over the resource split between training and quantized feedback as described in section VI-A. The curves for the MSE of CSIT obtained theoretically, by doing the simulations with actual constellations and the corresponding quantization bound for that constellation have been plotted in Fig. 6. Quantization bound gives the



(a)

(b)

Fig. 6. Mean-Square CSIT Errors with perfect CSIR: $T_{fb} = 20$ and $M = 4$ (a) QPSK and (b) 16-QAM. The novel hybrid scheme with QPSK performs better than the classical one from 7 to 25 dB of SNR, but 16-QAM outperforms both after 21 dB. Fixed split curves are for equal resource split between training and quantized feedback intervals.

quantization error when maximal ($T_{fb} - 1$) symbols are used for quantized feedback part. Hence, it gives the lower bound on the MSE of CSIT (performance upper bound) for that particular constellation. For comparison purpose, we have also plotted the MSE of CSIT for classical training scheme. This figure shows that from low to medium SNR values, the novel scheme with QPSK gives CSIT error below that of the classical training approach but 16-QAM is not attractive in this range due to many incorrect detection events. At high SNR values, hybrid scheme with QPSK suffers from performance degradation due to its bounded quantization error but 16-QAM behaves much better than the classical scheme. This dictates that larger constellations should be used for feedback with increasing SNR.

To demonstrate the value of optimal resource split, the curves of fixed split for both constellations have also been plotted in Fig. 6 where the feedback resource is equally split between training and quantized feedback ($T_a = T_q = T_{fb}/2$). We remark severe degradation for the fixed resource split compared to the optimal one as was pointed out in section V-B.

In Fig. 6, both for QPSK and 16-QAM, we have plotted the MSE of CSIT using our proposed iterative estimation and detection algorithms from section IV. One would expect the iterative estimation and detection algorithm (with ML detection) to perform better than the simplified iterative estimation and detection algorithm (which uses the simple LS detection), but extensive simulations show that the performance difference between the two algorithms is negligible. In all our simulations, both algorithms show very rapid convergence and they were always converging in second or third iteration. There were extremely rare instances (less than one in ten million) when convergence was not achieved in three iterations. We commented in section IV-A that the point of convergence depends upon the initial point. As in our system settings, this initial point is obtained through pure training based estimate and the system is operating in medium to large SNR regime, the quality of the initial estimate would be reasonably good. Further the presence of multiple receive antennas at the BS gives diversity and power gain in outage capacity (the key metric in the second phase of the proposed hybrid scheme). Hence the decoding of the feedback will normally fail only if all the channel coefficients are suffering through deep fades [30].

The closeness of the theoretical bounds and the system simulation curves in Fig. 6 shows the validity of the derived bounds and the analysis carried out in section V based upon the idea of outage rate. The difference is mainly due to beyond outage error events as both curves use bounded error for quantization. Without this upper bound of quantization error, the performance of hybrid scheme will improve further.

C. Discrete Constellations and Imperfect CSIR

As perfect CSIR assumption is too good to be true, in this subsection we remove this assumption and analyze how the MSE of CSIT with novel scheme behaves with imperfect CSIR.

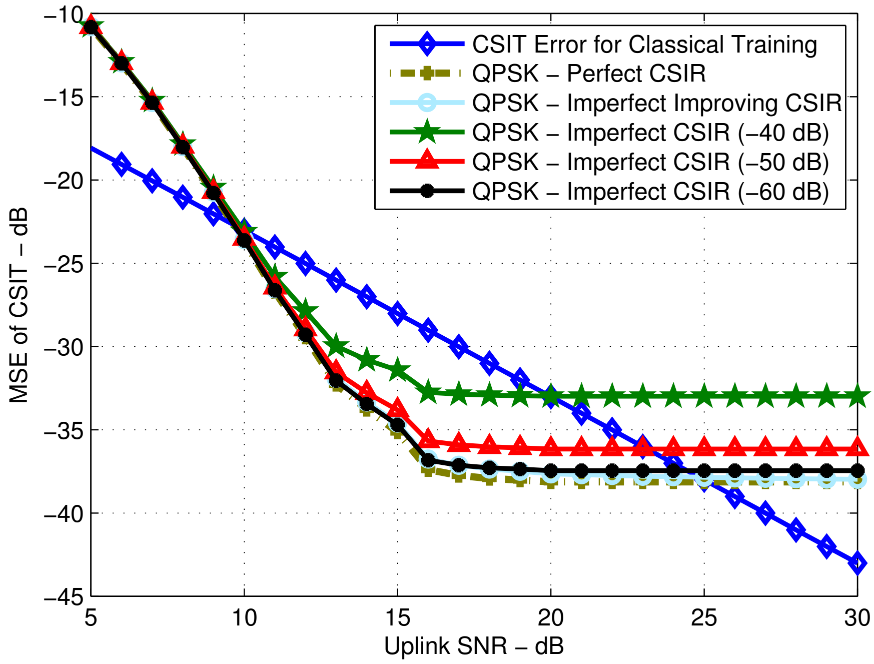
The curves, when quantized feedback is transmitted using QPSK and 16-QAM, have been plotted in Fig. 7. We have plotted these curves under two settings. First, when the CSIR quality varies and improves with the increase in UL SNR which is quite logical as, due to reciprocity, the link quality improves in both directions and the BS can surely pump more power as compared to a small hand-held mobile unit. For this case, we take the MSE of CSIR 30 dB less than the classical training only CSIT curve. The second scenario is when CSIR quality is held fixed independent of the UL SNR. For this, we plot the MSE of CSIT when the MSE of CSIR is kept fixed at -40, -50 and -60 dB. We believe this scenario to be of relatively less importance. We remark that when CSIR quality improves with UL SNR, hybrid approach performs very close to the perfect CSIR curve. For the other case when CSIR quality is kept fixed, it may become the performance limit of the MSE of CSIT (if not of proper quality).

D. Discrete Constellations and Coding

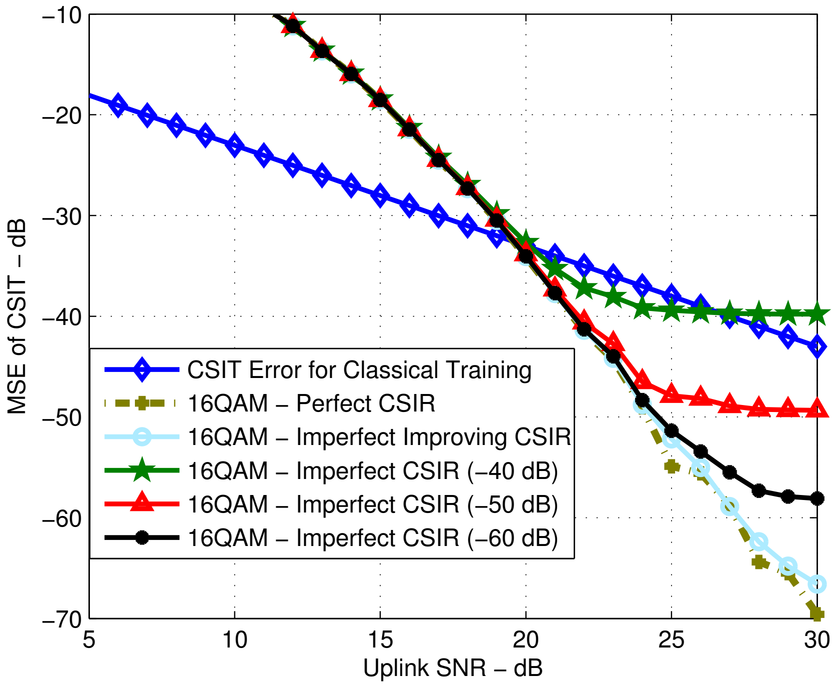
Now we plot the results of the MSE of CSIT when quantized feedback is sent using discrete constellations, the rate matching is performed using convolutional codes as explained in section VI-B and iterative estimation and detection algorithms are used at the BS. Here we assume only perfect CSIR. The code rates and the puncturing patterns need to be selected carefully. First of all, convolutional codes of all desired rates are not available. Secondly, although puncturing can help to achieve the desired rate, a random choice of puncturing pattern may destroy the code structure and hence ultimately its performance.

We plot the results obtained using three different codes ($1/2$ rate code, $2/3$ rate code and $3/4$ rate code) in Fig. 8. Fig. 4 has shown that the feedback rate should be below 4 bits per channel use till 28 dB so all of these codes have been used with 16-QAM (4 bits per channel use). Hence the number of actual information (feedback) bits b are 2, 2.67 and 3 per channel use for $1/2$, $2/3$ and $3/4$ rate code respectively. We optimize the split (getting T_a and T_q) using Theorem 2 for fixed b as in eq. (44). The codebook used is of size $2^{\lfloor bT_q \rfloor}$ and each single code index after coding/symbol mapping gives T_q sybmols of 16-QAM. For comparison purpose, the plot shows the MSE of CSIT obtained by using QPSK and 16-QAM constellations without any coding and through classical training scheme.

For $1/2$ rate code, the generator matrix is $[171\ 133]_8$ and trace back length is 30. It performs better than classical training from 16 to 23 dB of SNR but QPSK without any coding performs better than this curve. For $2/3$ rate code, the generator matrix is $[4\ 5\ 17; 7\ 4\ 2]_8$ with trace back length of 20. From



(a)



(b)

Fig. 7. Mean-Square CSIT Errors with Imperfect CSIR: $T_{fb} = 20$ and $M = 4$ (a) QPSK and (b) 16-QAM. For an imperfect CSIR of reasonable quality, the novel scheme performs much better than the classical scheme and the performance approaches to the perfect CSIR case for a good enough CSIR.

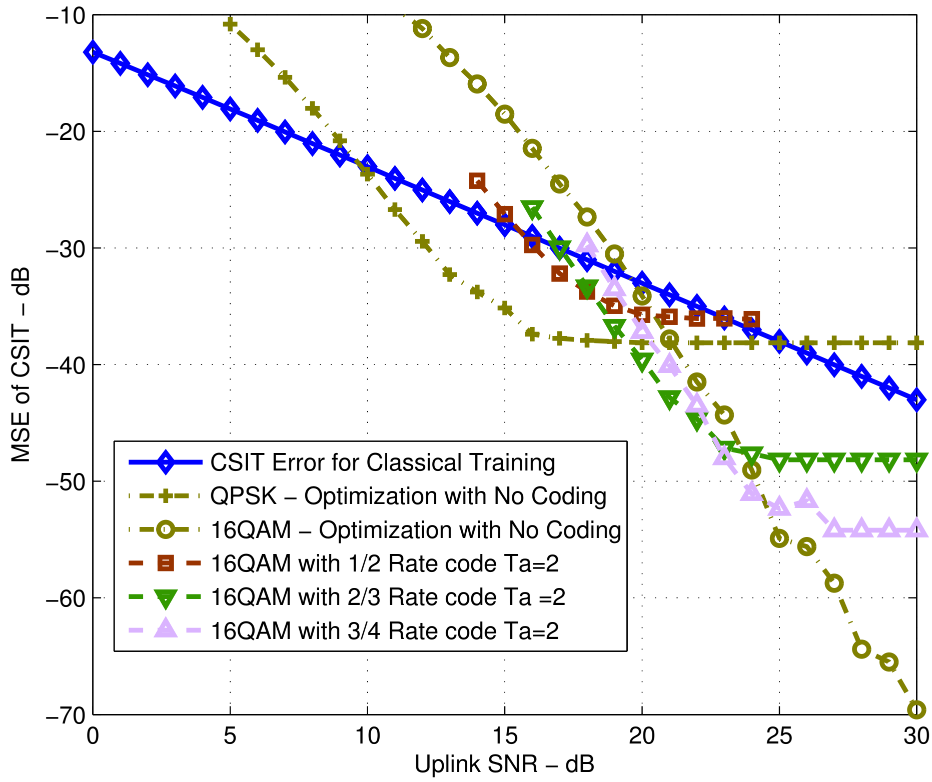


Fig. 8. Mean-Square CSIT Errors with Convolutional Coding: $T_{fb} = 20$ and $M = 4$. At certain SNR intervals, coding strategy performs better than no coding optimal resource split outcome.

17 dB onward, it performs better than classical training. It performs even better than 16-QAM (without coding) before 24 dB of SNR. For $3/4$ rate code, we use the $1/2$ rate base code (same as before) and use the puncturing pattern of $[111001]$ to get the final rate of $3/4$.

VIII. Concluding Remarks

Traditional CSIT acquisition in reciprocal systems relying exclusively on the use of training sequences ignores the shared knowledge of an identical channel between the BS and the user. We presented a novel approach of CSIT acquisition at the BS for the DL transmission in a reciprocal MIMO communication system combining the use of a training sequence together with quantized channel feedback. We characterized the optimal CSIT acquisition setup and proposed two iterative algorithms for the resulting joint estimation and detection problem and provided a convergence proof. The novel outage-rate based approach allows determining the optimal resource partitioning (between the training and the quantized feedback) and the feedback rate. We proposed two strategies to overcome the limitation of practical constellation availability with integer number of bits per channel use either by optimizing the resource split for a particular constellation or by the use of channel coding for rate matching. The novel combining scheme shows superior performance due to better exploitation of the reciprocity principle and the trade-off between the CSIT quality and the resource utilization improves significantly. It is further shown that with an imperfect CSIR of reasonable quality, performance gains comparable to the perfect CSIR case are achievable.

Multi-User Extension: The proposed novel scheme holds verbatim in the case of multiple users. In the first phase of "UL training", the users should use orthogonal training signals so that the BS gets an initial estimate of the channel. Then during the second "quantized feedback" phase, the UL channel should be used as MIMO-MAC. The optimization of resources remains however an open problem in this setting. In this scenario, the resource optimization will depend heavily upon the BS transmission strategy, e.g., the optimal resource split could be extremely different for TDMA or SDMA. The presence of more users in the system, larger than the BS transmit antennas, and subsequently required user scheduling would add an extra twist to this problem.

Users with Multiple Antennas: There are different ways to treat the fully general case of multiple users with multiple antennas where even a single user can be transmitted multiple streams. It adds an extra level of complexity to the open problem of multiple single-antenna users. For the users with multiple antennas, a simplifying strategy could be to do antenna combining as in [44] to minimize the quantization error. This scheme is promising as it reduces the feedback requirement by converting the MIMO channel into a vector channel and in a direction of minimal quantization error. Hence effectively it will become the multiple single-antenna user extension of our work.

ACKNOWLEDGMENTS

The research work carried out at Intel Mobile Communications leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) SACRA project (grant agreement n 249060).

EURECOM's research is partially supported by its industrial members: BMW Group, Swisscom, Cisco, ORANGE, SFR, ST Ericsson, Thales, Symantec, SAP, Monaco Telecom. The research of EURECOM is also supported in part by the EU projects ARTIST4G, CROWN, SACRA and WHERE2.

REFERENCES

[1] U. Salim, D. Gesbert, D. Slock, and Z. Beyaztas, “Hybrid pilot/quantization based feedback in multi-antenna TDD systems,” in *Proc. IEEE Global Communications Conference, Honolulu, HI, USA*, 2009.

[2] P. Viswanath and D. Tse, “Sum capacity of the multiple antenna Gaussian broadcast channel and uplink-downlink duality,” *IEEE Trans. on Information Theory*, vol. 49, pp. 1912–1921, August 2003.

[3] W. Yu and J. M. Cioffi, “Sum capacity of Gaussian vector broadcast channels,” *IEEE Trans. on Information Theory*, vol. 50, pp. 1875–1892, September 2004.

[4] D. Gesbert, M. Kountouris, J. R. W. Heath, C. B. Chae, and T. Salzer, “From single user to multiuser communications: Shifting the MIMO paradigm,” *IEEE Sig. Proc. Magazine*, 2007.

[5] P. Viswanath, D. N. C. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Transactions on Information Theory*, vol. 48, pp. 1277–1294, June 2002.

[6] M. Sharif and B. Hassibi, “On the capacity of MIMO broadcast channels with partial side information,” *IEEE Transactions on Information Theory*, vol. 51, pp. 506–522, February 2005.

[7] T. L. Marzetta, “How much training is required for multiuser MIMO?,” in *Proc. Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, November 2006, pp. 359–363.

[8] T. Marzetta and B. Hochwald, “Fast transfer of channel state information in wireless systems,” *IEEE Trans. on Signal Processing*, vol. 54, pp. 1268–1278, April 2006.

[9] U. Salim and D. Slock, “How much feedback is required for TDD multi-antenna broadcast channels with user selection?,” *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1–14, 2010.

[10] N. Jindal, “MIMO broadcast channels with finite rate feedback,” *IEEE Trans. on Information Theory*, vol. 52, pp. 5045–5060, November 2006.

[11] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, “What is the value of limited feedback for MIMO channels?,” *IEEE Communications Magazine*, vol. 42, pp. 54–59, October 2004.

[12] D. J. Love et al., “An overview of limited feedback in wireless communication systems,” *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 1341–1365, October 2008.

[13] T. Yoo, N. Jindal, and A. Goldsmith, “Multi-antenna downlink channels with limited feedback and user selection,” *IEEE Journal on Selected Areas in Communications*, vol. 24, September 2007.

[14] T. Pande, D. J. Love, and J. V. Krogmeier, “Reduced Feedback MIMO-OFDM Precoding and Antenna Selection,” *IEEE Trans. on signal processing*, vol. 55, pp. 2284–2293, May 2007.

[15] P. Xia and G. B. Giannakis, “Design and Analysis of Transmit-Beamforming based on Limited-Rate Feedback,” *IEEE Trans. on signal processing*, vol. 54, pp. 1853–1863, May 2006.

[16] P. Ding, D. J. Love, and M. D. Zoltowski, “Multiple Antenna Broadcast Channels with Shape Feedback and Limited Feedback,” *IEEE Trans. on signal processing*, vol. 55, pp. 3417–3428, July 2007.

[17] M. Guillaud, D. Slock, and R. Knopp, “A practical method for wireless channel reciprocity exploitation through relative calibration,” in *Proc. Intl. Symposium on Signal Processing and its Applications (ISSPA)*, Sydney, Australia, August 2005.

[18] F. Kaltenberger and M. Guillaud, “Exploitation of reciprocity in measured MIMO channels,” in *COST 2100, 9th Management Committee Meeting, TD(09)950, September 28-30, 2009, Vienna, Austria.*

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing - Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[20] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser mimo achievable rates with downlink training and channel state feedback," *IEEE Trans. on Information Theory*, vol. 56, pp. 2845–2866, June 2010.

[21] P. Tejera and W. Utschick, "Feedback of channel state information in wireless systems," in *Proc. International Conference on Communications*, 2007, pp. 908–913.

[22] K. R. Kumar and G. Caire, "Channel state feedback over the MIMO-MAC," in *Proc. IEEE Int. Symp. Information Theory*, 2009.

[23] M. M. Shanechi, R. Porat, and U. Erez, "Comparison of Practical Feedback Algorithms for Multiuser MIMO," in *Proc. IEEE Vehicular Technology Conference Spring*, April 2009.

[24] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Scheduling and pre-conditioning in multi-user MIMO TDD systems," in *Proc. IEEE International Conference on Communications, Beijing, China*, 2008, pp. 4100–4105.

[25] U. Salim and D. Slock, "Broadcast channel: Degrees of freedom with no CSIR," in *Proc. Allerton Conf. on Communication, Control, and Computing*, 2008.

[26] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Linear Precoding for Multi-User Multiple Antenna TDD Systems," *Arxiv preprint http://arXiv:0812.0621v2*.

[27] U. Salim and D. Slock, "How many users should inform the BS about their channel information?," in *Proc. International Symposium on Wireless Communication Systems*, 2009.

[28] C. Steger and A. Sabharwal, "Single-input two-way SIMO channel: diversity-multiplexing tradeoff with two-way training," *IEEE Trans. on Wireless Communications*, vol. 7, pp. 4877–4885, December 2008.

[29] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. on Information Theory*, vol. 49, pp. 1073–1096, May 2003.

[30] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge, U.K. Cambridge Univ. Press, 2005.

[31] A. J. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[32] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communications link in Rayleigh flat fading," *IEEE Trans. on Information Theory*, vol. 45, pp. 139–157, January 1999.

[33] LTE, *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation, Release 10, V.10.0.0*, 3GPP TS 36.211, 2011.

[34] D. Samardzija and N. B. Mandayam, "Unquantized and uncoded channel state information feedback in multiple antenna multiuser systems," *IEEE Transactions on Communications*, vol. 54, pp. 1335–1345, July 2006.

[35] A. Sutivong, M. Chiang, T. Cover, and Y.-H. Kim, "Channel Capacity and State Estimation for State-Dependent Gaussian Channels," *IEEE Trans. on Information Theory*, vol. 41, pp. 1486–1495, April 2005.

[36] A. D. Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. on Information Theory*, vol. 22, pp. 1–10, January 1976.

[37] H. Yamamoto, "Wyner-Ziv Theory for a General Function of the Correlated Sources," *IEEE Trans. on Information Theory*, vol. 28, pp. 803–807, September 1982.

[38] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array-Part 1: Algorithms," *IEEE Trans. on signal processing*, vol. 44, pp. 1184–1197, May 1996.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, United Kindom: Cambridge University Press, 1985.

[40] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Trans. on Vehicular Technology*, vol. 43, pp. 359–378, May 1994.

[41] E. Biglieri, J. Proakis, and S. Shamai, "Fading Channels: Information Theoretic and Communications Aspects," *IEEE Trans. on Information Theory*, vol. 44, pp. 2619–2692, October 1998.

[42] C. Au-Yeung and D. J. Love, "Optimization and Tradeoff Analysis of Two-Way Limited Feedback Beamforming Systems," *IEEE Trans. on Wireless Communications*, vol. 8, pp. 2570–2579, May 2009.

[43] C. Au-Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. on Wireless Communications*, vol. 6, pp. 458–462, February 2007.

[44] N. Jindal, "Antenna combining for the MIMO downlink channel," *IEEE Trans. on Wireless Communications*, vol. 10, pp. 3834–3844, October 2008.

Umer SALIM (S'06-M'10) received his PhD and MS both in electrical engineering with specialization in communication theory and signal processing from Eurecom (France) and Supelec (France) respectively. He is currently working at Intel Mobile Communications in the group of algorithm design where the main focus is on the design of efficient receivers for future wireless standards. Umer has several years of research experience in digital communications and signal processing and has published several papers in well-known conferences and journals. He co-authored a paper which got the best paper award in European Wireless Conference 2011. His mains areas of interest include signal processing techniques for multi-cell multi-user MIMO systems, novel and practical CSI feedback design techniques and analysis, information theoretic analysis of cognitive radio, and multi-user information theory in general.

David Gesbert (IEEE Fellow) is Professor and Head of the Mobile Communications Department, EURECOM, France, where he also heads the Communications Theory Group. He obtained the Ph.D degree from Ecole Nationale Superieure des Telecommunications, France, in 1997. From 1997 to 1999 he has been with the Information Systems Laboratory, Stanford University. In 1999, he was a founding engineer of Iospan Wireless Inc, San Jose, Ca.,a startup company pioneering MIMO-OFDM (now Intel). Between 2001 and 2003 he has been with the Department of Informatics, University of Oslo as an adjunct professor. D. Gesbert has published about 170 papers and several patents all in the area of signal processing, communications, and wireless networks.

D. Gesbert was a co-editor of several special issues on wireless networks and communications theory, for JSAC (2003, 2007, 2009), EURASIP Journal on Applied Signal Processing (2004, 2007), Wireless Communications Magazine (2006). He served on the IEEE Signal Processing for Communications Technical Committee, 2003-2008. He's an associate editor for IEEE Transactions on Wireless Communications and the EURASIP Journal on Wireless Communications and Networking. He authored or co-authored papers winning the 2004 IEEE Best Tutorial Paper Award (Communications Society) for a 2003 JSAC paper on MIMO systems, 2005 Best Paper (Young Author) Award for Signal Proc. Society journals, and the Best Paper Award for the 2004 ACM MSWiM workshop. He co-authored the book Space time wireless communications: From parameter estimation to MIMO systems, Cambridge Press, 2006.

Dirk T.M. Slock received an engineering degree from the University of Gent, Belgium in 1982. In 1984 he was awarded a Fulbright scholarship for Stanford University, USA, where he received the MSEE, MS in Statistics, and PhD in EE in 1986, 1989 and 1989 resp. While at Stanford, he developed new fast recursive least-squares algorithms for adaptive filtering. In 1989-91, he was a member of the research staff at the Philips Research Laboratory Belgium. In 1991, he joined EURECOM where he is now professor. At EURECOM, he teaches statistical signal processing (SSP) and signal processing techniques for wireless communications. His research interests include SSP for mobile communications (antenna arrays for (semi-blind) equalization/interference cancellation and spatial division multiple access (SDMA), space-time processing and coding, channel estimation, diversity analysis, information-theoretic capacity analysis, relaying, cognitive radio, geolocation), and SSP techniques for audio processing. He invented semi-blind channel estimation, the chip equalizer-correlator receiver used by 3G HSDPA mobile terminals, spatial multiplexing cyclic delay diversity (MIMO-CDD) now part of LTE, and his work led to the Single Antenna Interference Cancellation (SAIC) concept used in GSM terminals. Recent keywords are diversity-multiplexing tradeoff (DMT), MIMO interference channel, multi-cell, variational Bayesian techniques, large system analysis, user selection, audio source separation.

In 2000, he cofounded SigTone, a start-up developing music signal processing products. He has also been active as a consultant on xDSL, DVB-T and 3G systems. He is the (co)author of over 300 technical papers. He received one best journal paper award from IEEE-SP and one from EURASIP in 1992. He is the coauthor of two IEEE Globecom'98, one IEEE SIU'04 and one IEEE SPAWC'05 best student paper award, and a honorary mention (finalist in best student paper contest) at IEEE SSP'05, IWAENC'06 and IEEE Asilomar'06. He was an associate editor for the IEEE-SP Transactions in 1994-96 and the IEEE Signal Processing Letters in 2009-10. He is an editor for the EURASIP Journal on Advances in Signal Processing (JASP). He has also been a guest editor for JASP, IEEE-SP Sig. Proc. Mag. and for IEEE-COM JSAC. He was the General Chair of the IEEE-SP SPAWC'06 workshop. He is a Fellow of the IEEE.